Comparable Worth as Multiattribute Utility

Sarah Lichtenstein

Research Associate

Decision Research

A Branch of Perceptronics, Inc.

1201 Oak Street

Eugene, OR 97401

My topic today was recently characterized by the Reagan administration as "medieval." It's pay equity, specifically, comparable worth. I'll describe the methodology of comparable worth in detail, because comparable worth is the most exciting application of judgment and decision making ever. And so far, I seem to be one of only a few people in the world to recognize the link.

Let me start by defining comparable worth. To me, comparable worth is a pay policy whereby all the jobs in an organization are paid as an increasing function of their worth or value to the employer. It is a policy about pay within an organization. Since different employers have different needs or goals, it does not imply that all the janitors in the U.S. will be paid the same. It does say that janitors' pay will be systematically related to the pay of clerks in the same organization.

I'll use the word "bias" to refer to any systematic deviation from comparable worth. Bias means that the pay does not properly reflect the values of the employer. I'll use the word "compensable" to refer to job attributes the employer values in determining pay.

My definition says nothing about gender. But obviously the impetus for comparable worth is the conviction, held by many, that in our society female dominated jobs (that is, jobs held mostly by women, like secretary and nurse) are seriously underpaid relative to their worth. I share that conviction. I'm not going to talk about whether it's true, but assume it's true.

I'm also not going to talk about the legal status of comparable worth, except to note that, as I understand it, comparable worth is not legally mandated. The courts have not interpreted equal-pay-for-equal-work legislation as requiring comparable worth, and I don't expect them to.

Almost two years ago, the Oregon legislature passed a bill authorizing a comparable worth study of almost all state jobs (university professors was one of the excluded job classes). Last year the governor and legislative leaders appointed a 7-person Task Force to oversee the study. The study will be finished early next year, when a report will go to the legislature.

The Task Force appointed a 16-member Advisory Committee, mostly union and management people, but also three members representing the public. I am one of those three public members of the Advisory Committee. So I am getting a very quick education in comparable worth. Rather early on, I discovered, to my astonishment, that I am an expert in comparable worth. Indeed, as you will learn today, you are all experts. So this is an "I knew it all along" talk.

You'll be pleased to find out you're experts, because right now comparable worth is a hot topic. I don't know how many comparable worth studies there are, but surely more than 100 states, counties, and cities have recently completed, or are in the process of doing, a comparable worth study of their pay practices. And most of them are in desperate need of expert advice. We can find plenty of opportunities to give away our expertise -- or sell it.

I'd like to describe for you two rather different approaches to the measurement of comparable worth. The first might be called the "automated" approach; it requires little or no excercise of judgment. The literature on comparable worth calls this class of methods "policy capturing."

Start with a collection of items that describe job activities. One such collection of items is the PAQ, the Position Analysis Questionnaire. It has 178 items that refer to aspects of jobs such as whether the job

requires the use of written materials, or estimation of the speed of moving objects, or climbing, or personal contact with public customers. Most of the items are answered on a 6-point scale; a few are dichotomous.

Give the questionnaire to the job incumbents. Factor-analyze the results. Get job-factor scores for each job. Regress the job-factor scores against current wages. If gender discrimination is present in wages, it is most prevalent in female-dominated jobs, so this last step, regressing on wages, is done only for the male-dominated jobs. Voila! The resulting raw-score equation IS the new, comparable-worth-based pay policy, essentially untouched by human hands.

What can we say about this method? First, it is an approach that relies on the input; the items chosen for inclusion in the questionnaire. If few or no items relating to some compensable factor are included, that factor will not emerge from the factor analysis. If the items are worded in a biased way, that will affect the results. And so forth.

Second, it relies on present wages in a crucial way. If those wages do not reflect some compensable factor, that factor will not be included in the final pay policy. If you believe that the only effect of gender discrimination is that female-dominated jobs get lower pay, then this approach is okay. But if you believe that even wages now paid to male-dominated jobs contain bias, you won't like this method.

For example, you may believe that human-relations skills is a compensable factor. But current wages may not now be rewarding that skill. Or even if they do, maybe restricting the analysis to male-dominated jobs omits the jobs in which this factor shows much variance, by leaving out social workers, receptionists, and the like.

Shown on the first page of the handout are the results of such a study. (The unit for these numbers is $ per hour.) I was astounded when I saw these results. Of the five factors that had significant coefficients, three of the coefficients are NEGATIVE! That means that the more one's job requires a people orientation or physical demand or machine tending, the less it pays. Now, the authors were not entirely blind to this result. They did take note that using this pay policy, nurses -- one of the most notoriously underpaid jobs in America -- were found to be overpaid. The study was only a pilot project, to illustrate the method, not to establish a real pay policy. But the authors were not dissatisfied with their results because the results matched their definition of comparable worth, which was:

> This study defines comparable worth as existing when the empirical relationship between job content and wages is the same for male- and female-held jobs (p. 446).

I think this definition is a cop-out because it fails to recognize that the issue of fair wages is inherently an issue of values that cannot be avoided by reliance on elegant statistical techniques. Policy-capturing approaches will fail no matter how sophisticated the methodology, because they rely on current wages to establish relative values. The whole reason for doing a comparable worth study is that current wages are messed up; they are a fundamentally flawed criterion.

Let me turn now to the second approach to comparable worth, what is usually called the "point-factor" method. It is has also been called the a priori approach because the determination of relative values is made without reference to job-description data or current wages.

The world's largest corporation engaged in job evaluation is the Hay Group. They have been in the business since the 1940's, primarily in the

private sector; they now have offices in 20 American cities and 42 other cities worldwide. In recent years they have been doing more public-sector work, some, such as in Minnesota and San Jose, specifically oriented to comparable worth. They were hired by the state of Oregon to do its study, so I have had an opportunity to look into the nitty-gritty of their method.

The Hay method is the prototype for point-factor approaches. It is not only widely used but also widely copied. For example, the Willis Corporation, who did the Washington State study, uses a very similar method.

The Oregon study is huge. There are about 38,000 jobs covered by the study. These jobs are now organized into about 1200 job classes. In theory, all the jobs in a job class are essentially the same, but over the years the classification system has gotten out of whack, so everybody realized that the entire classification system had to be redone before the comparable-worth part of the study could begin.

The Task Force, working with the Hay Group, developed an open-ended, 13 page job questionnare, the same one for all jobs. It was given to all 38,000 job incumbents, who filled it out. Each questionnaire was also reviewed and, if necessary, added to by a supervisor. A new classification system then had to be developed. To do this, the completed questionnaires were put in one room, where someone had to sort them into about 1500 separate stacks, such that all the questionnaires within one stack described essentially the same job and no two stacks described the same job. Then a new questionnaire was filled out for each stack. This "composite" questionnaire was the quintessence of all the questionnaires in the stack. From this point on, the composites were used, not the original questionnaires.

Meanwhile, the Task Force, in hiring the Hay Group, had in essence purchased a pre-packaged comparable worth system. It has four factors, each of which is composed of two or three scales. The Task Force worked with Hay to make minor modifications in this system. The "human relations" scale was increased from three to four levels. The "working conditions" factor was extensively revised. Page 1 of the handout shows the names of the factors and the scales and shows the number of levels of each scale.

The next task was to translate the information contained in the composite questionnaires into the language of the Hay scales. The composites describe the jobs in terms of what the incumbents do, how often they do it, who supervises them and how often, what the most difficult parts of their jobs are and how they solve those difficulties, and so forth. Action language, not abstract language. In contrast, the Hay scales are described in quite general language, like:

> (from Managerial Know-how) Comprehensive integration and coordination of diversified activities and functions in a major management area.
> (from Specialized Techniques Know-how) Requires prolonged specialized training in job related courses in a formal curriculum or equivalent on-the-job experience and seasoning.
> (from Accountability Impact) Participating with others (except own subordinates and superiors), within or outside the organizational unit, in taking action.
> (from Working Conditions Work Demands) Having to perform in multiple, critical roles, or two or more incompatible roles.

Specially-trained raters, working in small groups, read each composite and rated that job on each of the 11 scales. The translation of these ratings to numbers was done using the Hay Guide Charts, one chart for each factor. One such Chart is shown in the handout. One chooses the relevant row and column and that choice determines a cell in the table. Each cell has two or three numbers in it. That's a kind of correction or adjustment -- or fudge factor. If you thought that the job was at or near the top of

its level on one or more of the scales, you would pick the largest of the numbers in the cell.

Each chart, covering one factor, produces one number. These numbers are then added together to get the total number of points associated with the job, with one exception. The Problem-solving factor is considered to be mediated by Know-how. So the entry in the Problem-solving chart is a percentage. You multiply this percentage by the Know-how score previously found in order to get the Problem-solving score. For example, if the Know-how score were 132 and the cell you chose in the Problem-solving chart said 33%, then the Problem-solving score would be 44.

It looks kind of complicated, but in essence, what the Hay Guide Chart system is, it's a multi-attribute utility model or what Keeney and Raiffa call a multi-attribute value model.

It was the responsibility of the Task Force to determine if this canned MAUT correctly captured or expressed the values of the employer, the state of Oregon. I believe they were considerably hampered in that determination by the way the MAUT was presented, via the Guide Charts. I thought it would be helpful to present the same information in an equation that expressed the MAUT model. So I uncovered -- detected-- that underlying model. What I developed is not just an approximation of the Hay Guide Charts. It really is the underlying model, although the Hay people seemed unaware of it. I started by assigning what I'll call rating numbers, r's, to each level of each scale. I assigned them in such a way that the <u>sum</u> of the rating numbers for the worst, the least compensable, job in each chart would be equal to 0. I also discovered that I could not reproduce the Hay Guide Charts unless I incremented some rating numbers by one for each next level, and others by two or three. The rating numbers I

derived are shown on page 3 of the handout, along with the numbers I
assigned to the correction factors. Given this indexing system, the
underlying MAUT model for the Hay Guide Charts is shown in the equation
below.

Isn't that a thing of beauty? The combination rule isn't additive.
It isn't multiplicative. It's a weird beast: exponential within factors
and additive between factors. Notice also that the first three scales
appear twice in the equation, once in the Know-how factor and again in the
Problem-solving factor. That's because the latter is used to multiply the
former to get the Problem-solving points.

Let me try to explain why this model is (partially) exponential. The
Hay group relies on Weber's Law. They say that when a rater says that two
jobs differ by just one level, the jobs are really 15% different (well,
14.87% gives a better fit, but that's because of rounding in the charts).
They consider 15% to be about a jnd in job evaluation. I don't buy this
explanation. I agree that judgments of, say, lengths of lines increase
less and less as the real, true, objectively-measured length increases
linearly. But there is no real, true, objectively-measured criterion for,
say, the degree of human relations skills required by a particular job.
The judgment of a knowledgeable and carefully trained rater or group of
raters is the best assessment we have; I see absolutely no reason to
rescale it.

The worst thing about this model is that you just can't tell by
looking at it whether it does or does not express your values. I've been
trying to come up with a way of characterizing the relative importance of
each scale in this model. Now, not all MAUT advocates agree that relative
importance is a meaningful concept. Keeney and Raiffa say:

>"If we have assessed [the weight for scale Y as] .75 and [the weight for scale Z as] .25, we <u>cannot</u> say that Y is three times as important as Z. In fact, we cannot conclude that attribute Y is more important than Z. Going one step further, it is not clear how we would precisely define the concept that one attribute is more important than another" (p. 272; emphasis in original).

But I would contend that in this situation, where the model has not been built from the judgments of the present users but has been presented to them as a package, it is essential to provide those users with some measure of relative importance. That turned out to be a toughie. An intuitively appealing first step would be to ask, "How does the total score change as each scale, in turn, is increased by one level?" That gets us nowhere. The answer depends not only on which level of the chosen scale you start at, but also on which level you're holding constant for all the other scales.

Well, how about "How does the total score change as we move from the lowest level to the highest level of each scale?" That, too, depends on what is assumed about the other scales. A further problem with this lowest-to-highest approach is that the actual ratings using the Hay system are going to be highly intercorrelated. I predict that the <u>lowest</u> intercorrelation among the first eight scales, that is, not counting Working Conditions, will be around .75, and that many will exceed .90. There are five possible reasons for these intercorrelations. First, there may be legitimate, conceptually distinct compensable factors that really are correlated in the population of jobs being evaluated. Second, halo effects. The raters rate one composite on all 11 scales before going on to the next composite. That's a situation ripe for halo effects. Third, some of the Hay scales are conceptually very similar, or even identical. Page 4 of the handout shows bits of the wording attached to one scale in the

Problem-solving factor and one scale in the Accountability factor. They sure sound the same to me. Fourth, the training done by the Hay Group discourages raters from coming up with "profiles" of ratings that the Hay Group believes are unrealistic or impossible. The raters are, in effect, discouraged from making ratings that would lower the correlations among scales. Fifth, of course, is the fact that the first three scales appear in both the first and the second factor.

Anyhow, these intercorrelations make an approach to relative importance based on going from worst to best on one scale, holding all else constant, unattractive.

Since this is a hierarchical MAUT model, it seems reasonable to approach the issue of relative importance hierarchically. Start with each factor separately. Look only at the scales within that factor, and satisfy yourself that those scales correctly reflect your relative values. For this, it is important to note both the number of levels in each scale and whether the rating numbers increase by one, two, or three steps when moving from one level to the next. These two aspects are nicely captured in the rating range that I have shown in the last column of page 3 in the handout. That's just the difference between the highest rating and the lowest rating for each scale. I initially thought that these rating ranges could serve well to measure relative importance within a factor. Within a factor, the rating numbers are nicely additive. So increasing one scale by one rating number will have the same effect as increasing some other scale by one rating number. Easy to interpret. But there is a further complication. The raters will not use all of the levels. Or they will use some levels so rarely that they hardly count.

For example, observers in Oregon tell me that the raters have interpreted the lowest level of human relations skills so narrowly that only people who never have to speak to <u>anyone</u> in the performance of their job are given this rating. That decreases the effective range of the scale.

My best idea so far is to use as a measure of relative importance within factors the standard deviation of the rating numbers. I really like that measure -- but -- it can't be computed until after all the ratings have been done. So it is useless to the Task Force in deciding ahead of time whether the model captures their values. And let me lay on you one further problem. Suppose I had the data and were about to compute the change in the sum of the ratings for a factor if each scale in turn were increased by one standard deviation. Should I hold all the other scales constant at their means or should I increment <u>them</u> a bit, too, as a function of how much each of them is correlated with the chosen scale? I really have no idea what the answer to that question is, and I wish someone here would think about it for me.

Now let's talk about relative importance of factors. The constants in the equation are relevant, here. They give the score for the worst job for each factor. But they don't tell the whole story, because there are huge range differences among the factors.

For example, the highest score on Accountability is 1840; the highest on Know-how is less, 1400, even though Know-how has an apparent weight four times as great. So I think we are driven to measure relative importance by looking at the geometric standard deviation of the scores for each factor.

The model has another peculiarity. Generally speaking, an irregular profile of ratings earns more points than an average profile. This

tendency is so strong that one can even find a set of profiles, starting with an average for each factor and producing a series of more and more irregularity, such that even though the total scores from a purely additive model or a purely exponential model would go down across the series, the total scores from this mixed model go up. So is this a compensatory model or isn't it?

There are a few comments one can make with reasonable confidence about relative importance in this model. For example, human relations skills has less importance than managerial skills. And working conditions accounts for zilch (the Hay Group's experience has been primarily with upper-level or "exempt" employees; their bias towards management skills is fairly evident). In the end, however, I advised my friends on the Oregon Task Force that, for the most part, I didn't know what relative values were being expressed in the Hay System and I didn't believe anyone else did, either.

I wasn't very happy with that "expert" opinion. However, I also told them that because of the enormous intercorrelations to be expected among the scales, bias in relative importance probably wouldn't make much of a difference. Halo conquers all.

There is one kind of halo that worries me. I'll dub it "gender halo." Suppose the raters know that a job is female-dominated. Suppose they assume, without realizing their assumption, that female-dominated jobs just couldn't require the levels of responsibility, accountability, and what-have-you that male-dominated jobs require. That sort of halo would produce large and systematic bias in the system. And I'm not sure whether training could overcome it.

I think that, as judgment and decision researchers, we could make the biggest contribution by shedding light on the models being used today and what those models say about values. But also, as experimental psychologists, we can say quite a bit about potential sources of bias in the whole comparable worth process. We are, I hope, less likely to assume that factor analysis and regression are value free, and we have some knowledge of where the bodies are hidden in these techniques. We can access the judgment literature, the attitude measurement literature, and the social cognition literature to suggest subtle sources of bias. We know that, to some extent, language determines or constrains thought, so we can take a careful look at the words used in the questionnaires and guide charts. For example, here's an item from the PAQ:

> How important is setting up or adjusting equipment (setting up a lathe or drill press, adjusting an engine carburetor, etc.)?

This item was rewritten for a comparable worth study to read:

> How important is setting up or adjusting equipment (attaching devices to patients, setting up a lathe or drill press, adjusting office equipment)?

Before the rewording, respondents might, without realizing it, answer this question as if it were asking the degree to which the job required traditionally male-only equipment skills. The rewording makes this interpretation less likely.

Although I think we have a lot to offer, I am somewhat pessimistic about avoiding bias. It is easy to fall back on unexamined habits, and because our culture's values about women and work are changing rapidly right now, these old, automatic attitudes may not represent the values we now espouse. My favorite example of this is about cleanliness and dirtiness. It comes from a discussion by Helen Remick, who notes that male

tree trimmers for the city of Denver, interviewed for a television report
on the efforts of nurses in that city to raise their wages, repeatedly said
that they thought tree trimmers deserved a higher salary than nurses
because their work was more "difficult," "dangerous," and "dirty." She
comments:

> Their choice of terms illustrates how cultural values
> invade our concepts of jobs -- and efforts to conduct bias-free
> job evaluations. It would seem fair to conclude that the tree
> trimmers are referring to the physical difficulty of climbing
> trees and ladders, the danger of physical labor at heights and
> with certain machinery, and the dirt of outdoors work. They,
> and many others, do not see the difficulty of work in intensive
> care units, the danger of dealing with disease and psychotic
> patients, or the dirt of vomit. I find their use of "dirty"
> especially interesting. Apparently, to most men and women
> alike, dirty jobs are those where no attempt is made to keep the
> work environment clean, and the dirt shows, under your
> fingernails, after work hours; axle grease is dirty. Many
> nurses I have talked to see their job as clean, in part because
> of the constant effort to make the environment sterile, in spite
> of their exposure to vomit, urine, feces, blood, pus, dead
> people, disease, and so on.

Okay, back to the point-factor method. We've bought this model, we've
done the ratings, and we now have a total Hay score for each job class.
What we have to do next is set a pay policy. How much does a job with 762
Hay points get paid?

One could do a MAUT analysis of this problem, eliciting values for the
shape, slope, and elevation of the pay policy line. Or one could leave it
entirely up to union-management negotiations or legislative whim. But for
practical or political reasons, regression of some sort will probably be
used. The regression of Hay points against wages should first be tested
for curvilinearity. God knows what kind of curvature that crazy semi-
exponentiality will produce. I also advised against a dog-leg line. The
Hay Group is fond of dog-legs. I've shown a dog-leg in the handout. I

think an abrupt change in slope is psychologically unsound. Better to use a curved line or rescale one of the variables to straighten the plot out.

There is also considerable sentiment for using only the male-dominated jobs in the regression to establish a pay policy. The reasoning is that because of years of discrimination against women, one would not expect the average wages of female-dominated jobs to represent fair compensation. Remember that the intention is to pay all employees according to the pay policy. Why should both men and women be penalized for past discrimination against women?

I want to make clear that this use of regression, to set a pay policy, does not suffer from the faults of using regression in the policy-capturing method. Here, regression is not being used to determine what the compensable attributes are or to determine the relative weights of those attributes. The only assumption being made here is that current wages for male-dominated jobs reasonably reflect the employer's values concerning the shape, slope, and elevation of the pay policy.

But there is an ethical peculiarity of using this approach. Wouldn't you agree that the slope of the pay policy is a union-management issue? I mean, typically, unions want the ratio of the highest-paid job to the lowest-paid job to be relatively small, while management wants it to be relatively large. But notice that any random error in current wages will flatten the slope. It seems strange to me to say that the larger the past goofs in pay and classification, the more the union position will prevail. Likewise, random error in Hay points will flatten the slope. Does it then follow that unions should encourage the raters to be sloppy in their ratings?

Because of this ethical peculiarity, I would advocate a key-job approach to setting the pay policy. The key-job method has a well-established history in the pay-evaluation field. Select just a few job classes, such that they nicely span the range of Hay points. Each key job class is a class that can be unambiguously and easily described, that has a ready market (I mean, the job is typically filled from outside the organization and there is a moderate amount of turnover in the job in that region), and, for reasons I gave earlier, that is male-dominated. Now do a careful market survey of these key jobs in the region. Use those data for establishing a pay policy, either by regression or by fairing in a line. This method is time consuming. And it is open to bias when the key jobs are selected. But it does have the advantage of being sensitive to market forces for those jobs that are market controlled.

There are two additional steps in any comparable worth study that I will mention only for completeness. First, the pay policy must be expanded to include variations in pay for compensable attributes not found in a whole job class, such as merit or seniority pay, or bonuses for night-shift work. Second, the implementation of a new pay policy can be a nightmare. How much will it cost, and will the legislature fund it? How fast must the underpaid jobs be raised to avoid a law suit? What should one do with the overpaid jobs? Cuts in pay are usually out of the question. Even freezes in pay might be. And unions are often opposed to paying new employees less than the incumbents (while one waits for the incumbents to retire). Finally, to what extent will unions oppose comparable worth because it limits the scope of their negotiations?

I'd like to turn now to my view of the criticisms that are being made about the whole idea of comparable worth.

1. They say, "It can't be done." This one fascinates me, because in some cases this is said by the very people who earn their living doing job evaluations. I really don't see that calling a job evaluation by the new name of comparable worth makes it impossible to do.

2. They say, "It defies the market." If you believe that in the U.S. there is an unconstrained job market, comparable worth is not for you. But there is a lot of evidence against that. I read with interest that there was testimony in the Denver nurses case that hospital administrators in the city met every year for the explicit purpose of agreeing on nurses' pay, thereby constraining the market.

3. They say, "It will be too expensive." Well, it will cost money. But the whole process will be spread out over several years. And cost is, I think, a weak argument to justify injustice.

4. They say, "It will lead to economic disruption." Not if introduced gradually, as it surely will be. And besides, this is an old and tired argument. Child labor laws did not cause our economy to fail, as was once predicted. Nor did minimum wage laws, though dire effects were predicted then, too.

5. They say, "The whole problem is now being taken care of with affirmative action." Sure.

And 6., my favorite: "There is no theoretical basis for comparable worth. You can't compare apples and oranges." These critics seem unaware of multiattribute utility theory, and I think we ought to bring it to their attention. What other social program has not only a theoretical basis, but, for gosh sakes, an _axiomatic_ base?

There _is_ one interesting theoretical issue for comparable worth. It arises when the employer is not a single person whose values can be assessed, but an organization like a state. If the state is viewed as a

collection of people with conflicting values, Social Choice Theory says that there is no way to reasonably represent those values in a single, majority-choice ordering. That's because Social Choice Theory rejects the possibility of interpersonal value comparisons such as "My preference for tea over coffee is greater than your preference for gin over vodka".

Keeney and Raiffa, in their book on MAUT, are willing to make interpersonal value comparisons. This assumption allows the calculation of a group value function. I feel pretty comfortable about this assumption for comparable worth. We do not have to make value judgments about jobs, like "which is more important to the state, fixing the roads or collecting taxes?" Instead, the attributes are qualities that are not job-specific. So the comparisons we are making are more like, "Which attribute should have more effect on pay, management skills or technical skills?" While such judgments are not easy, I think they can be argued about on fixed, shared value scales, each going from the worst level of an attribute set at 0 to the best level of an attribute set at 1.

You see, you're all instant experts. I haven't talked about anything new today. You all knew that random error makes regression slopes flatter. And that values can be measured. And that hierarchical models can be useful in organizing one's beliefs. I urge you all to offer your expertise when the occasion arises. You're needed. I think it's fair to characterize the current literature on comparable worth, by and large, as shockingly naive from a technical standpoint. I plan to do something about that, myself, as soon as I can get the funding.

Since you're experts, I need your advice. What can one tell users of the Hay system about the relative values it expresses? Are there any

sensible conclusions one can draw from that exponential-additive model? If so, please let me know.

The last page of the handout gives a few references to the comparable worth literature.

Finally, some words about data. I am reasonably optimistic that I will be able to get my hands on all the composite job questionnaires and on the ratings given to them. Oh, the research possibilities just scream out. Everything from "Are there differences in ratings made by male and female raters?" to "How much differently would Oregon state employees be paid if Oregon had used an additive model?" It will take some months before the data are available. I can hardly wait.

HANDOUT -- PAGE 1

Pierson, Koziara, and Johannesson (in Remick, 1984 and in Public
Personnel Management, 1983) developed the following raw-score
regression equation to predict wages:

Factors	Coefficients
(Constant)	5.57
Cognitive Judgment	.57
People-orientation	-.43
Complexity	.47
Physical demands	-.40
Machine tending	-.21
Working conditions	.04
Word & paper processing	.00
Reading & listening	.06

Factors and scales of the Hay Guide Charts used in Oregon:

Name	Number of levels
Know-How	
Learned, Spec., Prac.[a]	8
Managerial Know-How	4
Human Relations Skills	4
Problem Solving	
Thinking Environment	8
Thinking Challenge	5
Accountability	
Freedom to Act	8
Impact	4
Magnitude	5
Working Conditions	
Sensory/Muscular Effort	4
Physical Environment/Hazards	4
Work Demands	3

[a]Learned Disciplines, Specialized Techniques, and Practical
Procedures

STATE OF OREGON
COMPARABLE WORTH STUDY
JULY 1984

GUIDE HAY CHART
ACCOUNTABILITY
© HAY ASSOCIATES 1984

DEFINITION: Accountability is the answerability for an action and for the consequences thereof. It is the measured effect of the job on end results. It has three dimensions in the following order of importance:

- **Freedom to Act** - the degree of personal or procedural control and guidance as defined in the left-hand column below.
- **Job Impact on End Results** - as defined at upper right.
- **Magnitude** - the effect of the job relative to the overall results of the STATE, which may be reflected by impact on the achievement of significant objectives leading toward the overall mission of the STATE. (Dollar values are to be used as guidelines only in assessing magnitude. Dollar values are in constant dollars, 1965 Base.)

IMPACT OF JOB ON END RESULTS

- **REMOTE:** Informational, recording, routine, or specific, task-oriented duties or services for use by others in taking action in relation to some important end result. *(INDIRECT)*
- **CONTRIBUTORY:** Interpretive, advisory, or facilitating services for use by others in taking action. *(INDIRECT)*
- **SHARED:** Participating with others (except own subordinates and superiors), within or outside the organizational unit, in taking action. *(DIRECT)*
- **PRIMARY:** Controlling impact on end results, where any shared accountability of others is secondary. *(DIRECT)*

Magnitude descriptions:
- **(1) VERY SMALL OR INDETERMINATE:** Major job effect is an results with an indeterminate or very small relationship to the missions of the total corporation, in relation to the total objectives of the agency, the position impacts on a very small or indeterminate portion of the objectives.
- **(2) SMALL:** Major job effect is an results of an element of that organization directed toward a part of STATE mission. In relation to the total objectives of the STATE the position impacts on the small portion of the objective. This may encompass a portion of income, expenditure, or controlled resources and assets.
- **(3) MEDIUM:** Major job effect is an results of a medium element typically directed toward real STATE missions. In relation to the total objectives of the organization, the position impacts on a significant medium portion of the objectives. This may encompass a medium portion of the income, expenditure or controlled resources and assets.
- **(4) LARGE:** Major job effect is an results of a large element of the STATE directed toward one of the missions of the STATE. In relation to the total objectives of the STATE, the position impacts on a large portion of the objective. This may encompass a large portion of the income, expenditure, or controlled resources and assets.
- **(5) VERY LARGE:** Major job effect is an overall results of the STATE and all portion of its missions in relation to the total objectives of the STATE, the position impacts on a majority or all the objective. It may encompass a substantial portion of income, expenditure, and controlled resources and assets.

AMI for use with 1983 dollars is ___.

Freedom to Act	(1) Very Small / Indeterminate (Under $100M)				(2) Small ($100M - $1MM)				(3) Medium ($1MM - $10MM)				(4) Large ($10MM - $100MM)				(5) Very Large ($100MM - $1MMM)			
	R	C	S	P	R	C	S	P	R	C	S	P	R	C	S	P	R	C	S	P
A. PRESCRIBED 1	10	14	19	25	14	19	25	33	19	25	33	43	25	33	43	57	33	43	57	76
A 2	12	16	22	29	16	22	29	38	22	29	38	50	29	38	50	66	38	50	66	87
A 3	14	19	25	33	19	25	33	43	25	33	43	57	33	43	57	76	43	57	76	100
B. CONTROLLED 1	16	22	29	38	22	29	38	50	29	38	50	66	38	50	66	87	50	66	87	115
B 2	19	25	33	43	25	33	43	57	33	43	57	76	43	57	76	100	57	76	100	132
B 3	22	29	38	50	29	38	50	66	38	50	66	87	50	66	87	115	66	87	115	152
C. STANDARDIZED 1	25	33	43	57	33	43	57	76	43	57	76	100	57	76	100	132	76	100	132	175
C 2	29	38	50	66	38	50	66	87	50	66	87	115	66	87	115	152	87	115	152	200
C 3	33	43	57	76	43	57	76	100	57	76	100	132	76	100	132	175	100	132	175	230
D. GENERALLY REGULATED 1	38	50	66	87	50	66	87	115	66	87	115	152	87	115	152	200	115	152	200	264
D 2	43	57	76	100	57	76	100	132	76	100	132	175	100	132	175	230	132	175	230	304
D 3	50	66	87	115	66	87	115	152	87	115	152	200	115	152	200	264	152	200	264	350
E. DIRECTED 1	57	76	100	132	76	100	132	175	100	132	175	230	132	175	230	304	175	230	304	400
E 2	66	87	115	152	87	115	152	200	115	152	200	264	152	200	264	350	200	264	350	460
E 3	76	100	132	175	100	132	175	230	132	175	230	304	175	230	304	400	230	304	400	528
F. GUIDANCE 1	87	115	152	200	115	152	200	264	152	200	264	350	200	264	350	460	264	350	460	608
F 2	100	132	175	230	132	175	230	304	175	230	304	400	230	304	400	528	304	400	528	700
F 3	115	152	200	264	152	200	264	350	200	264	350	460	264	350	460	608	350	460	608	800
G. GENERAL GUIDANCE 1	132	175	230	304	175	230	304	400	230	304	400	528	304	400	528	700	400	528	700	920
G 2	152	200	264	350	200	264	350	460	264	350	460	608	350	460	608	800	460	608	800	1056
G 3	175	230	304	400	230	304	400	528	304	400	528	700	400	528	700	920	528	700	920	1216
H. GOVERNOR/CHIEF JUSTICE 1	200	264	350	460	264	350	460	608	350	460	608	800	460	608	800	1056	608	800	1056	1400
H 2	230	304	400	528	304	400	528	700	400	528	700	920	528	700	920	1216	700	920	1216	1600
H 3	264	350	460	608	350	460	608	800	460	608	800	1056	608	800	1056	1400	800	1056	1400	1840

Freedom to Act row descriptions:

- **A. PRESCRIBED** — These jobs are subject to: Direct and detailed instructions / Close supervision
- **B. CONTROLLED** — These jobs are subject to: Instructions and established work routines / Close supervision
- **C. STANDARDIZED** — These jobs are subject, wholly or in part, to: Standardized practices and procedures / General work instructions / Supervision of progress and results
- **D. GENERALLY REGULATED** — These jobs are subject, wholly or in part, to: Practices and procedures covered by precedents or well-defined policy / Supervisory review
- **E. DIRECTED** — These jobs, by their nature or size, are subject to: Broad practice and procedures covered by functional precedents and policies / Managerial direction
- **F. GUIDANCE** — These jobs are inherently subject only to broad policy and general management guidance.
- **G. GENERAL GUIDANCE** — These jobs, by reason of their nature or size, independent complexity, and high degree of effect on State operations are subject only to Guidance from the Governor's office.
- **H. GOVERNOR/CHIEF JUSTICE** — These jobs are subject only to the limitations of the State Constitution as it pertains to the Executive or Judicial Branch.

HANDOUT -- PAGE 3

Model for the Hay System used in Oregon

Let j index the scales, j = 1 to 11.
Let r stand for rating number, as assigned below
Let c stand for the correction factor, as assigned below

Scale Name	Scale No. (j)	Rating No. (r)	Correction Factor	Rating Range
Know-How			$c_k=-1,0,+1$	
Learned, Spec., Prac.	1	1, 3, 5,...15		14
Managerial Know-How	2	0, 2, 4, 6		6
Human Relations Skills	3	0, 1, 2, 3		3
Problem Solving			$c_p=0,+1$	
Thinking Environment	4	0, 1, 2,...7		7
Thinking Challenge	5	0, 2, 4, 6, 8		8
Accountability			$c_a=-1,0,+1$	
Freedom to Act	6	1, 4, 7,...22		21
Impact	7	0, 2, 4, 6		6
Magnitude	8	0, 2, 4, 6, 8		8
Working Conditions			$c_w=-1,0,+1$	
Sensory/Muscular Effort	9	1, 3, 5, 7		6
Physical Envir./Hazards	10	0, 2, 4, 6		6
Work Demands	11	0, 2, 4		4

Then Total Hay Points (H) may be found with the equation:

$$H = 43.5\,(1.1487)^{\sum_{j=1}^{3} r_j + c_k} + 4.7\,(1.1487)^{\sum_{j=1}^{5} r_j + c_k + c_p}$$

$$+ 10.9\,(1.1487)^{\sum_{j=6}^{8} r_j + c_a} + 4.1\,(1.1487)^{\sum_{j=9}^{11} r_j + c_w}$$

HANDOUT -- PAGE 4

Wording from the Thinking Environment and Freedom to Act Scales.

Thinking Environment   Freedom to Act

A. Rigidly Structured. Very detailed and precisely defined rules and instructions and/or continually present assistance. Instructions usually specify in detail the sequence and timing of the tasks to be undertaken with little latitude permitted to consider alternative procedures.

A. Prescribed. These jobs are subject to direct and detailed instructions and close supervision.

B. Routine. Detailed standard practices and/or immediately available assistance or examples. Instructions usually provide latitude to consider variations in sequence of procedures based on situations encountered with the work setting.

B. Controlled. These jobs are subject to instructions and established work routines and close supervision.
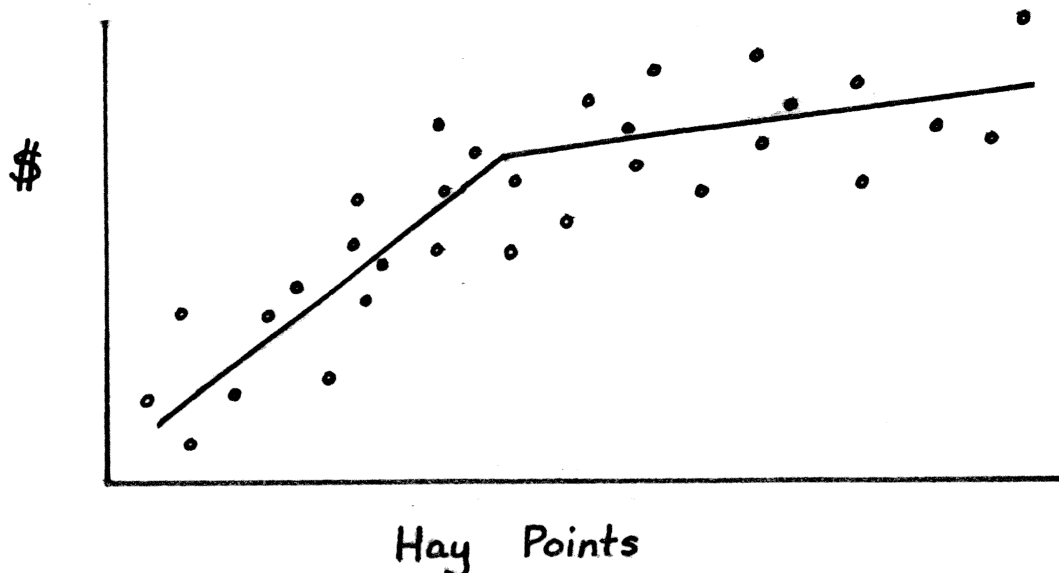
. . . . .

. . . . .

F. Broadly Defined. Broad policies and objectives, under general direction. Organizational policies and functional goals shape the scope of thinking freedom.

F. Guidance. These jobs are inherently subject only to broad policy and general management guidance.

. . . . .

. . . . .

Example of a dog-leg pay policy line:

Hay Points

Selected References

Livernash, E. R. (Ed.). (1984). <u>Comparable worth: Issues and alternatives</u>, second edition. Washington, DC: Equal Employment Advisory Council.

This is the best "anti" volume.

Reichenberg, N. E. (Ed.). (1983). Comparable worth [Special issue]. <u>Public Personnel Management</u>, <u>12</u> (4).

Whole issue of this journal devoted to the topic. Neither overwhelmingly "pro" nor "anti."

Remick, H. (Ed..). (1984). <u>Comparable worth and wage discrimination: Technical possibilities and political realities</u>. Philadelphia: Temple University Press.

This is the best "pro" volume.

Treiman, D. J. (1979). <u>Job evaluation: An analytic review</u>. Washington, DC: National Academy of Sciences.

Treiman, D. J., & Hartmann, H. I. (Eds.). (1981). <u>Women, work, and wages: Equal pay for jobs of equal value</u>. Washington, DC: National Academy Press.

This pair of reports shows considerable sophistication about policy capturing methods but less understanding of point-factor methods.